

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 10011093 A

(43) Date of publication of application: 16.01.98

(51) Int. Cl.

G10L 9/14

G10L 9/18

(21) Application number: 08164840

(71) Applicant: NEC CORP

(22) Date of filing: 25.06.96

(72) Inventor: OZAWA KAZUNORI

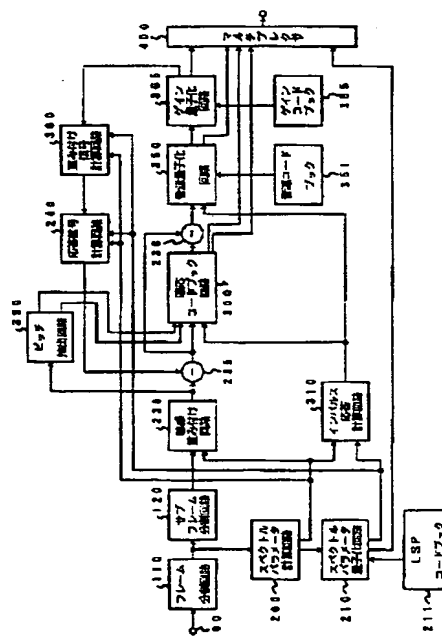
(54) SIGNAL ENCODING DEVICE

(57) Abstract:

PROBLEM TO BE SOLVED: To obtain excellent sound quality for voices from more than one speaker and musical sound signals generated by more than one musical instrument by calculating a sound source signal after finding delays from input signals.

SOLUTION: A spectrum parameter calculating circuit 200 performs calculation in predetermined degree of a spectrum parameter by providing a window which is longer than subframe length for a voice signal of at least one subframe and segmenting a voice. A pitch extracting circuit 390 finds delays and gains corresponding to a pitch cycle by using the output of an auditory weighting circuit 230. Then a sound source quantizing circuit 350 performs vector quantization for a sound source signal by using a sound source code book 351. The output of a subtracter 236 and the output of an impulse response calculating circuit 310 are used to retrieve a sound source code vector $c_i(n)$ from the sound source code book 351. Further, the index of the selected sound source vector is outputted to a multiplexer 400.

COPYRIGHT: (C)1998,JPO



BEST AVAILABLE COPY

特開平10-11093

(43) 公開日 平成10年(1998) 1月16日

(51) Int.Cl.⁶

G10L 9/14

9/18

識別記号

庁内整理番号

FI

G10L 9/14

9/18

技術表示箇所

G

J

E

審査請求 有 請求項の数 8 OL (全 20 頁)

(21) 出願番号

特願平8-164840

(22) 出願日

平成8年(1996) 6月25日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 小澤 一範

東京都港区芝五丁目7番1号 日本電気株式会社社内

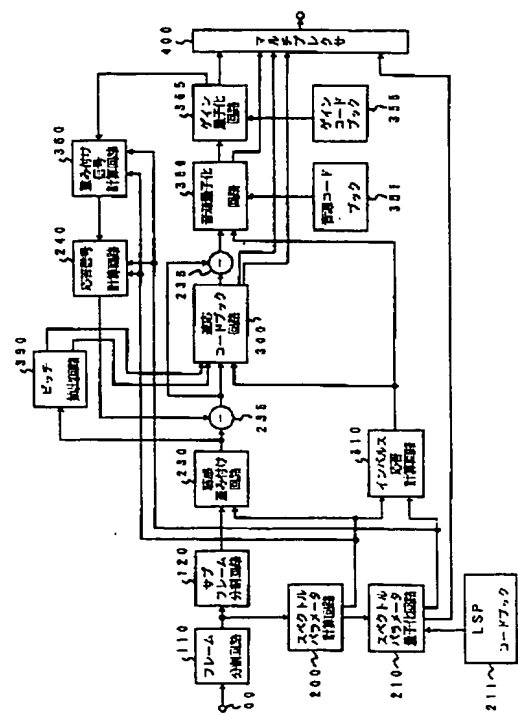
(74) 代理人 弁理士 後藤 洋介 (外2名)

(54) 【発明の名称】 信号符号化装置

(57) 【要約】

【課題】 音声ならず音楽信号に対して低ビットレートでも良好な音質の得られる信号符号化装置の提供。

【解決手段】 入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部200と、前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求めこれらの処理を繰り返しあらかじめ定められた個数の遅延を求めて出力する遅延計算部390と、前記複数個の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部350を有することを特徴とする。



【特許請求の範囲】

【請求項1】 入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求め、これら第1及び第2の遅延を少なくとも含む、あらかじめ定められた複数の遅延を求めて出力する遅延計算部と、前記複数の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置。

【請求項2】 前記信号の音源信号が振幅が非零の複数のパルスから構成されることを特徴とする請求項1記載の信号符号化装置。

【請求項3】 入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から特徴量を抽出してモードを判別するモード判別部と、あらかじめ定められたモードにおいて前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求め、これら第1及び第2の遅延を少なくとも含む、あらかじめ定められた複数の遅延を求めて出力する遅延計算部と、前記複数の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置。

【請求項4】 前記信号の音源信号が振幅が非零の複数のパルスから構成されることを特徴とする請求項3記載の信号符号化装置。

【請求項5】 入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め第1のピッチ予測利得を計算し、前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求めこれらの処理を繰り返すピッチ予測部と、前記ピッチ予測利得があらかじめ定められた条件を満たすかどうかを判別する判別部と、前記ピッチ予測利得があらかじめ定められた条件を満たさない場合に前記遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置。

【請求項6】 前記信号の音源信号が振幅が非零の複数のパルスから構成されることを特徴とする請求項5記載の信号符号化装置。

【請求項7】 入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から特徴量を抽出してモードを判別するモード判別部と、あらかじめ定められたモードにおいて前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め第1のピッチ予測利得を計算

し、前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求めこれらの処理を繰り返すピッチ予測部と、前記ピッチ予測利得があらかじめ定められた条件を満たすかどうかを判別する判別部と、前記ピッチ予測利得があらかじめ定められた条件を満たさない場合は前記遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置。

【請求項8】 前記信号の音源信号が振幅が非零の複数のパルスから構成されることを特徴とする請求項7記載の信号符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声や音楽信号を低いビットレートで高品質に符号化するための信号符号化装置に関するものである。

【0002】

【従来の技術】音声信号を高効率に符号化する方式としては、例えば、M. Schroeder and B. Atalによる“Code-excited line ar prediction: High quality speech at very low bit rates” (Proc. ICASSP, pp. 937-940, 1985年)と題した論文(文献1)や、Kleijnらによる“Improved speech quality and efficeintvect or quantization in SELP” (Proc. ICASSP, pp. 155-158, 1988年)と題した論文(文献2)などに記載されているCELP (Code Excited Linear Predictive Coding) が知られている。この従来例では、送信側では、フレーム毎(例えば20ms)に音声信号から線形予測(LPC)分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。フレームをさらにサブフレーム(例えば5ms)に分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータ(ピッチ周期に対応する遅延パラメータとゲインパラメータ)を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。ピッチ予測して求めた音源信号に対して、予め定められた種類の雑音信号からなる音源コードブック(ベクトル量子化コードブック)から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化するように行う。そして、選択されたコードベクトルの種類を表すインデックスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせて伝送する。受信側の説明は省略す

る。

【0003】

【発明が解決しようとする課題】前記従来法では、音源コードブックから最適な音源コードベクトルを選択するのに多大な演算量を要するという問題があった。これは、文献1や2の方法では、音源コードベクトルを選択するのに、各コードベクトルに対して一旦フィルタリングもしくは畳み込み演算を行ない、この演算をコードブックに格納されているコードベクトルの個数だけ繰り返すことに起因する。例えば、コードブックのビット数がBビットで、次元数がNのときは、フィルタリングあるいは畳み込み演算のときのフィルタあるいはインパルス応答長をKとすると、演算量は1秒当たり、 $N \times K \times 2B \times 8000 / N$ だけ必要となる。一例として、 $B=10$ 、 $N=40$ 、 $K=10$ とすると、1秒当たり81,920,000回の演算が必要となり、極めて膨大であるという問題点があった。また、この問題点は、入力信号の帯域が電話帯域よりも広く、標準化周波数が高くなるほど、深刻であった。

【0004】音源コードブック探索に必要な演算量を低減する方法として、種々のものが提案されている。例えば、ACELP (Algebraic Code Excited Linear Prediction) 方式が提案されている。これは、例えば、C. Laflammeらによる“16 kbps wideband speech coding technique based on algebraic CELP”と題した論文 (Proc. ICASSP, pp. 13-16, 1991) (文献3) 等を参照することができる。文献3の方法によれば、音源信号を複数個のパルスで表し、各パルスの位置をあらかじめ定められたビット数で表し伝送する。ここで、各パルスの振幅は+1.0もしくは-1.0に限定されているため、パルス探索の演算量を大幅に低減化できる。

【0005】しかし、以上述べたいずれの手法も、ピッチが一つの音声に対しては比較的良好な音質が得られるものの、会議などの用途で、複数話者の声や、複数種の楽器が混在するためにピッチが複数個含まれる音楽信号に対しては、低いビットレートでは甚だしく音質が劣化していた。

【0006】本発明の目的は、上述の問題を解決し、ビットレートが低い場合にも、広帯域の音声のみならず音楽信号に対しても、比較的少ない演算量で音質の劣化の少ない信号符号化方式を提供することにある。

【0007】

【課題を解決するための手段】本発明の第1の態様によれば、入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め前記ピッチ予測残差信号から第2の

ピッチ周期に対応する第2の遅延を求め、これら第1及び第2の遅延を少なくとも含むあらかじめ定められた複数個の遅延を求めて出力する遅延計算部と、前記複数個の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置が得られる。

【0008】本発明の第2の態様によれば、前記信号の音源信号が振幅が非零の複数個のパルスから構成されることを特徴とする信号符号化装置が得られる。

【0009】本発明の第3の態様によれば、入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から特徴量を抽出してモードを判別するモード判別部と、あらかじめ定められたモードにおいて前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求め、これら第1及び第2の遅延を少なくとも含むあらかじめ定められた複数個の遅延を求めて出力する遅延計算部と、前記複数個の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置が得られる。

【0010】本発明の第4の態様によれば、前記信号の音源信号が振幅が非零の複数個のパルスから構成されることを特徴とする信号符号化装置が得られる。

【0011】本発明の第5の態様によれば、入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め第1のピッチ予測利得を計算し、前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求めこれらの処理を繰り返すピッチ予測部と、前記ピッチ予測利得があらかじめ定められた条件を満たすかどうかを判別する判別部と、前記ピッチ予測利得があらかじめ定められた条件を満たさない場合に前記遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置が得られる。

【0012】本発明の第6の態様によれば、前記信号の音源信号が振幅が非零の複数個のパルスから構成されることを特徴とする信号符号化装置が得られる。

【0013】本発明の第7の態様によれば、入力信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記入力信号から特徴量を抽出してモードを判別するモード判別部と、あらかじめ定められたモードにおいて前記入力信号から第1のピッチ周期に対応する第1の遅延を求めピッチ予測残差信号を求め第1のピッチ予測利得を計算し、前記ピッチ予測残差信号から第2のピッチ周期に対応する第2の遅延を求めこれらの処理を繰り返すピッチ予測部と、前記ピッチ予測

利得があらかじめ定められた条件を満たすかどうかを判別する判別部と、前記ピッチ予測利得があらかじめ定められた条件を満たさない場合は前記遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有することを特徴とする信号符号化装置が得られる。

【0014】本発明の第8の態様によれば、前記信号の音源信号が振幅が非零の複数のパルスから構成されることを特徴とする信号符号化装置が得られる。

【0015】本発明の第1の態様では、入力信号からピッチ周期に相当する遅延の計算とピッチ予測残差信号の計算を繰り返し、あらかじめ定められた個数の遅延を求め、前記複数の遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する。

【0016】本発明の第2の態様では、第1の態様において、音源信号が個数Mの振幅が非零のパルス列から構成され、パルスの振幅と位置を求めることにより、音源信号を量子化する。

【0017】本発明の第3の態様では、入力信号から特徴量を抽出してモードを判別し、あらかじめ定められたモードにおいてのみ、第1の発明と同一の動作を行なう。

【0018】本発明の第4の態様では、第3の態様において、音源信号が個数Mの振幅が非零のパルス列から構成され、パルスの振幅と位置を求めることにより、音源信号を量子化する。

【0019】本発明の第5の態様では、入力信号からピッチ周期に相当する遅延の計算とピッチ予測残差信号の計算とピッチ予測利得の計算を繰り返し、前記ピッチ予測利得があらかじめ定められた条件を満たすかどうかを判別し前記ピッチ予測利得があらかじめ定められた条件を満たさない場合に前記遅延を用いてピッチ予測した残差信号に対して音源信号を求めて量子化して出力する音源量子化部を有する。

【0020】本発明の第6の態様では、第5の態様において、音源信号が個数Mの振幅が非零のパルス列から構成され、パルスの振幅と位置を求めることにより、音源信号を量子化する。

【0021】本発明の第7の態様では、入力信号から特徴量を抽出してモードを判別し、あらかじめ定められたモードにおいてのみ、第5の態様と同一の動作を行う。

【0022】本発明の第8の態様では、第7の態様において、音源信号が個数Mの振幅が非零のパルス列から構成され、パルスの振幅と位置を求めることにより、音源信号を量子化する。

【0023】

【発明の実施の形態】次に本発明の実施例について図面を参照して説明する。

【0024】図1は本発明による音声符号化装置の一実施例を示すブロック図である。

【0-0-2-5】図において、入力端子1-0-0から音声信号を入力し、フレーム分割回路110では音声信号をフレーム（例えば10ms）毎に分割し、サブフレーム分割回路120では、フレームの音声信号をフレームよりも短いサブフレーム（例えば5ms）に分割する。

【0026】スペクトルパラメータ計算回路200では、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓（例えば24ms）をかけて音声を取り出してスペクトルパラメータをあらかじめ定められた次数（例えば $P=10$ 次）計算する。ここでスペクトルパラメータの計算には、周知のLPC分析や、Burg分析等を用いることができる。ここでは、Burg分析を用いることとする。Burg分析の詳細については、中溝著による“信号解析とシステム同定”と題した単行本（コロナ社1988年刊）の82～87頁（文献4）等に記載されているので説明は略する。さらにスペクトルパラメータ計算部では、Burg法により計算された線形予測係数 α_i （ $i=1, \dots, 10$ ）を量子化や補間に適したLSPパラメータに変換する。ここで、線形予測係数からLSPへの変換は、菅村他による“線スペクトル対（LSP）音声分析合成方式による音声情報圧縮”と題した論文（電子通信学会論文誌、J64-A、pp. 599-606、1981年）（文献5）を参照することができる。例えば、第2サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第1サブフレームのLSPを直線補間により求めて、第1サブフレームのLSPを逆変換して線形予測係数に戻し、第1、2サブフレームの線形予測係数 α_{i1} （ $i=1, \dots, 10$ 、 $l=1, \dots, 2$ ）を聴感重み付け回路230に出力する。また、第2サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0027】スペクトルパラメータ量子化回路210では、あらかじめ定められたサブフレームのLSPパラメータを効率的に量子化する。量子化法として、ベクトル量子化を用いるものとし、第2サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。具体的な方法は例えば、特開平4-171500号公報（特願平2-297600号）（文献6）や特開平4-36300号公報（特願平3-261925号）（文献7）や、特開平5-6199号公報（特願平3-155049号）（文献8）や、T. Nomura et al., による“LSP Coding Using VQ-AVQ With Interpolation in 4.075 kbps M-LCELP Speech Coder”と題した論文（Proc. Mobile Multimedia Communications, pp. B. 2. 5, 1993）（文献9）等を参照できる。

【0028】LSPコードブック211を用いて、下記の数式1で表わされる歪みを最小化するコードベクトルを選択して出力する。

$$D_j = \sum_{i=1}^P W(i) [LSP(i) - QLSP(i)_j]^2 \quad (1)$$

ここで、LSP(i)、QLSP(i)_j、W(i)はそれぞれ、量子化前のi次目のLSP、量子化後のj番目の結果、重み係数である。

【0030】また、スペクトルパラメータ量子化回路210では、第2サブフレームで量子化したLSPパラメータをもとに、第1サブフレームのLSPパラメータを復元する。ここでは、現フレームの第2サブフレームの量子化LSPパラメータと1つ過去のフレームの第2サブフレームの量子化LSPを直線補間して、第1サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後に、直線補間により第1サブフレームのLSPを復元できる。

【0031】以上により復元した第1サブフレームのLSPと第2サブフレームの量子化LSPをサブフレーム毎に線形予測係数 α'_i (i=1, ..., 10)に変換し、インパルス応答計算回路310へ出力する。また、第2サブフレームの量子化LSPのコードベクトルを表

$$x_z(n) = d(n) - \sum_{i=1}^P \alpha_i d(n-i) + \sum_{i=1}^P \alpha_i \gamma^i y(n-i) + \sum_{i=1}^P \alpha'_i \gamma^i x_z(n-i) \quad (2)$$

但し、 $n-i \leq 0$ のときは

$$y(n-i) = p(N+(n-i)) \quad (3)$$

$$x_z(n-i) = s_w(N+(n-i)) \quad (4)$$

ここでNはサブフレーム長を示す。 γ は、聴感重み付け量を制御する重み係数であり、下記の式(6)と同一の値である。 $s_w(n)$ 、 $p(n)$ は、それぞれ、重み付け信号計算回路の出力信号、後述の式(6)における右辺第1項のフィルタの分母の項の出力信号をそれぞれ示す。

【0035】ピッチ抽出回路390は、聴感重み付け回路230の出力を用いてピッチ周期に対応する遅延とゲインを複数個求める。以下ではこの個数を2とする。ピ

$$E_T = \sum_{n=0}^{N_1} x_v^2(n)$$

$$- \left[\sum_{n=0}^{N-1} x_v(n) x_v(n-T_1) \right]^2 / \left[\sum_{n=0}^{N-1} x_v^2(n-T_1) \right]$$

(5)

【0029】

【数1】

すインデックスをマルチプレクサ400に出力する。

【0032】聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に量子化前の線形予測係数 α_i (i=1, ..., P)を入力し、前記文献1にもとづき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号を出力する。

【0033】応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 α_i を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α'_i をサブフレーム毎に入力し、保存されているフィルタメモリの値を用いて、入力信号を零 $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算器235へ出力する。ここで、応答信号 $x_z(n)$ は下記の数式2で表される。

【0034】

【数2】

ッチ抽出回路390の構成を図2に示す。

【0036】図2において、端子391から聴感重み付け信号 $x_w(n)$ を入力する。第1の遅延、ゲイン計算回路392は、下記の数式3によって表わされる式

(5)を最小化するように第1の遅延 T_1 を求め、下記の数式4によって表わされる式(6)から第1のゲイン β_1 を求める。

【0037】

【数3】

$$\beta_1 = \frac{\sum_{n=0}^{N-1} x_w(n) x_w(n-T_1)}{\sum_{n=0}^{N-1} x_w^2(n-T_1)} \quad (6)$$

さらに、次式に従い第1のピッチ予測信号 $y_1(n)$ を 【0039】
求め減算器394に出力する。

$$y_1(n) = \beta_1 x_w(n-T_1) \quad (7)$$

減算器394は、次式により、第1のピッチ予測残差信 【0040】
号 $e_1(n)$ を求める。

$$e_1(n) = x_w(n) - y_1(n) \quad (8)$$

第2の遅延、ゲイン計算回路393は、 $e_1(n)$ から 10
第2の遅延 $T_2(n)$ 、ゲイン β_2 を求める。これら
は、式(5)、(6)において、 $x_w(n)$ のかわりに
 $e_1(n)$ を用いれば良い。

【0041】 T_2 、 T_1 は、それぞれ、端子397、3
98から出力される。

【0043】

【数5】

【0042】インパルス応答計算回路310は、 z 変換

$$H_v(z) = \frac{1 - \sum_{i=1}^P a_i z^{-i}}{1 - \sum_{i=1}^P a_i \tau^i z^{-i}} \cdot \frac{1}{1 - \sum_{i=1}^P a'_i \tau^i z^{-i}} \quad (9)$$

適応コードブック回路300は、 T_1 の近傍のサンプル
において、下記の数式6を最小化する遅延 T_{c1} を計算す
る。なお、ここでは、適応コードブックの次数を1とす

る。

【0044】

【数6】

$$D_{T_{c1}} = \sum_{n=0}^{N-1} x_w'^2(n) - \left[\sum_{n=0}^{N-1} x_w'(n) y_w(n-T_{c1}) \right]^2 / \left[\sum_{n=0}^{N-1} y_w^2(n-T_{c1}) \right] \quad (10)$$

ここで、

$$y_w(n-T_{c1}) = v_1(n-T_{c1}) * h_w(n) \quad (11)$$

であり、記号*は畳み込み演算を表す。

【0046】

【0045】ゲイン β_{c1} を下記の数式7に従い求める。

【数7】

$$\beta_{c1} = \frac{\sum_{n=0}^{N-1} x_w'(n) y_w(n-T_{c1})}{\sum_{n=0}^{N-1} y_w^2(n-T_{c1})} \quad (12)$$

ここで、女性音や子供の声に対して、遅延の抽出精度を
向上させるために、遅延を整数サンプルではなく、小数
サンプル値で求めてもよい。具体的な方法は、例えば、
P. Kroonらによる、“Pitch predictors with high temporal r
esolution”と題した論文(Proc. ICA
SSP, pp. 661-664, 1990年) (文献1

1)等を参照することができる。

【0047】同様の手法により、 T_2 の近傍のサンプル
において、遅延 T_{c2} とゲイン β_{c2} を探索する。

【0048】次に、ピッチ予測信号を計算し、音源量子
化回路350に出力する。

【0049】

$$q_{w1}(n) = \beta_{c1} v(n-T_{c1}) * h_w(n) + \beta_{c2} v(n-T_{c2}) * h_w(n) \quad (13)$$

遅延 T_{c1} 、 T_{c2} はマルチプレクサ400に出力される。

50 【0050】減算器236は、次式の計算を行ない、減

算結果を音源量子化回路350へ出力する。

【0051】

$$z_w(n) = x'_w(n) - q_w(n) \quad (14)$$

音源量子化回路350では、音源信号を、音源コードブック351を用いてベクトル量子化する。減算器236の出力と、インパルス応答計算回路310の出力を用いて、下記の数式8を最小化するように、音源コードブック

ク351から音源コードベクトル $c_j(n)$ を探索する。

【0052】

【数8】

$$D_j = \sum_{n=0}^{N-1} z_w(n)^2 - \frac{\left[\sum_{n=0}^{N-1} d(n) c_j(n) \right]^2}{\sum_{n=0}^{N-1} s_{wj}(n)^2} \quad (15)$$

$$d(n) = \sum_{i=n}^{N-1} z_w(i) h_w(i-n), \quad n=0, \dots, N-1 \quad (16)$$

$$s_{wj}(n) = c_j(n) * h_w(n) \quad (17)$$

ここで、記号*は、畳み込み演算を示す。

【0053】選択された音源コードベクトルのインデックスは、マルチプレクサ400に出力される。

【0054】ゲイン量子化回路365は、ゲインコードブック355からゲインコードベクトルを読みだし、選択された音源コードベクトルに対して、下記の数式9を

$$D_t = \sum_{n=0}^{N-1} [z_w(n) - G'_t s_{vj}(n)]^2 \quad (18)$$

ここで、 G'_t は、ゲインコードブック355に格納されたゲインコードブックにおける t 番目のコードベクトルである。

【0057】選択されたゲインコードベクトルを表すインデックスをマルチプレクサ400に出力する。

$$v(n) = g(n) + G'_t c_j(n) \quad (19)$$

$v(n)$ を適応コードブック回路300に出力する。

【0060】次に、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて下記の数式10によ

20 最小化するゲインコードベクトルを選択する。

【0055】ここでは、音源コードベクトルのゲインをベクトル量子化する例について示す。

【0056】

【数9】

30 【0058】重み付け信号計算回路360は、それぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読みだし、まず下式にもとづき駆動音源信号 $v(n)$ を求める。

【0059】

り、応答信号 $s_w(n)$ をサブフレーム毎に計算し、応答信号計算回路240へ出力する。

【0061】

【数10】

$$s_w(n) = v(n) - \sum_{l=1}^P a_l v(n-l)_{40}$$

$$+ \sum_{l=1}^P a_l \gamma^l p(n-l) + \sum_{l=1}^P a'_l \gamma^l s_w(n-l)$$

(20)

以上により、第1の発明に対応する実施例の説明を終える。

【0062】図3は、第2の実施例の構成を示すブロッ

ク図である。図において、図1と異なるのは、音源量子化回路500、振幅コードブック540、ゲイン量子化回路550、ゲインコードブック560である。

【0063】音源量子化回路500は、M個の振幅が非零のパルス列の位置と振幅を計算する。

【0064】音源量子化回路500の構成を示すブロック図を図4に示す。図4において、相関係数計算回路510は、端子501、502からそれぞれ、 $x_w(n)$ 、 $h_w(n)$ を入力し、下記の数式11及び数

$$d(n) = \sum_{l=n}^{N-1} z(l) h_w(i-n), \quad n=0, \dots, N-1 \quad (21)$$

【0066】

10 【数12】

$$\phi(p, q) = \sum_{n=\max(p, q)}^{N-1} h_w(n-p) h_w(n-q), \quad p, q=0, \dots, N-1 \quad (22)$$

位置計算回路520は、あらかじめ定められた個数Mの非零の振幅のパルスの位置を計算する。これには、文献3と同様に、各パルス毎に、あらかじめ定められた位置の候補について、次式を最大化するパルスの位置を求める。

【0067】例えば、位置の候補の例は、サブフレーム長を $N=40$ 、パルスの個数を $M=5$ とすると、下記の表1のように表せる。

【0068】

【表1】

$$D = C_k^2 / E_k$$

ここで、 C_k 及び E_k は下記の数式13及び14により表わされる。

$$C_k = \sum_{k=1}^M \text{sgn}(k) d(m_k)$$

30

【0071】

$$E_k = \sum_{k=1}^M \text{sgn}(k)^2 \phi(m_k, m_k)$$

$$+ 2 \sum_{k=1}^{M-1} \sum_{l=k+1}^M \text{sgn}(k) \text{sgn}(l) \phi(m_k, m_l) \quad (25)$$

ここで、 m_k は、 k 番目のパルスの位置を示す $\text{sgn}(k)$ は k 番目のパルスの極性である。

【0072】M個のパルスの位置は振幅量子化回路530に出力される。

$$C_j^2 / E_j$$

ここで、 C_j 及び E_j は下記の数式15及び16により表わされる。

$$C_j = \sum_{k=1}^N g'_k d(m_k)$$

【0076】

式12に従い、2種の相関係数 $d(n)$ 、 ϕ を計算し、位置計算回路520、振幅量子化回路530に出力する。

【0065】

【数11】

0.	5.	10.	15.	20.	25.	30.	35
1.	6.	11.	16.	21.	26.	31.	36
2.	7.	12.	17.	22.	27.	32.	37
3.	8.	13.	18.	23.	28.	33.	38
4.	9.	14.	19.	24.	29.	34.	39

各パルスについて、位置の候補を調べ、次式を最大化する位置を選択する。

【0069】

(23)

【0070】

【数13】

(24)

【数14】

【0073】振幅量子化回路530は、パルスの振幅を振幅コードブック540を用いて量子化する。次式を最大化する振幅コードベクトルを選択する。

【0074】

(26)

【0075】

【数15】

(27)

【数16】

$$E_j = -\sum_{k=1}^N g'_{kj} \phi - (m_k, m_k) + 2 \sum_{k=1}^{N-1} \sum_{l=k+1}^N g'_{kl} g'_{lj} \phi - (m_k, m_l)$$

(28)

ここで、 g'_{kj} は、 i 番目の振幅コードベクトルにおける k 番目のパルスの振幅を示す。

【0077】なお、パルスの振幅を量子化するための振幅コードブックを、音声信号を用いてあらかじめ学習して格納しておくこともできる。コードブックの学習法は、例えば、Linde氏らによる“An algorithm for vector quantization design,”と題した論文(IEEE Trans. Commun., pp. 84-95, January, 1980)(文献12)等を参照できる。

$$D_t = \sum_{n=0}^{N-1} [z_w(n) - G'_t \sum_{k=1}^M g'_{kj} h_w(n-m_k)]^2 \quad (29)$$

重み付け信号計算回路570は、それぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読みだし、まず下式にもとづき駆動音源信号 v

$$v(n) = g(n) + G'_t \sum_{k=1}^M g'_{kj} h_w(n-m_k) \quad (30)$$

$v(n)$ を適応コードブック回路300に出力する。

【0082】次に、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて下記の数式18によ

$$s_w(n) = v(n) - \sum_{i=1}^P a_i v(n-i) + \sum_{i=1}^P a_i \tau^i p(n-i) - 30 \sum_{i=1}^P a'_i \tau^i s_w(n-i)$$

(31)

図5は第3の実施例の構成を示すブロック図である。

【0084】モード判別回路900は、聴感重み付け回路230からフレーム単位で聴感重み付け信号を受取り、モード判別情報をピッチ抽出回路600、マルチプレクサ400に出力する。

【0085】ここでは、モード判別に、現在のフレーム

$$G = 10 \log_{10} [1/L \sum_{i=1}^L (P_i/E_i^{40})] \quad (32)$$

ここで、 L はフレームに含まれるサブフレームの個数である。 P_i 、 E_i はそれぞれ、 i 番目のサブフレームでの音声パワー、ピッチ予測誤差パワーを示し、下記の数

$$P_i = \sum_{n=0}^{N-1} x_{wi}^2(n) \quad (33)$$

【0088】

【0078】振幅コードベクトルのインデックスと位置の情報は、それぞれ、端子503、504から出力される。

【0079】ゲイン量子化回路550は、ゲインコードブック560を用いてパルスのゲインを量子化する。下記の数式17を最小化するようなゲインコードベクトルを選択し、インデックスをマルチプレクサ400へ出力する。

【0080】

【数17】

(n) を求める。

【0081】

り、応答信号 $s_w(n)$ をサブフレーム毎に計算し、応答信号計算回路240へ出力する。

【0083】

【数18】

の特徴量を用いる。特徴量としては、例えば、フレームで平均したピッチ予測ゲインを用いる。ピッチ予測ゲインの計算は、例えば下記の数式19を用いる。

【0086】

【数19】

式20及び21により表わされる。

【0087】

【数20】

【数21】

$$E_1 - P_1 - \frac{\sum_{n=0}^{N-1} x_w(n) x_w(n-T)}{\sum_{n=0}^{N-1} x_w^2(n-T)}$$

(34)

ここで、Tは予測ゲインを最大化する最適遅延である。

【0089】フレーム平均ピッチ予測ゲインGをあらかじめ定められた複数個のしきい値と比較して複数種類のモードに分類する。モードの個数としては、例えば4を用いることができる。

【0090】ピッチ抽出回路600は、モード判別情報を入力し、あらかじめ定められたモードの場合に図2と同一の処理を行ない、複数個の遅延を出力する。それ以外のモードでは、遅延の出力は行なわない。

【0091】図6は、第4の実施例の構成を示すブロック図である。図5におけるモード判別回路900を図3に付加し、ピッチ抽出回路600を用いたものであるので、説明は省略する。

$$G_1 = 10 \log_{10} \frac{\sum_{n=0}^{N-1} x_w^2(n)}{\sum_{n=0}^{N-1} [x_w(n) - \beta_1 x_w(n-T_1)]^2} \quad (dB)$$

(35)

判別回路730は、ピッチ予測ゲインG₁があらかじめ定められたしきい値よりも大きい場合は第2の遅延、ゲイン計算回路に対して処理を継続させる。

【0096】G₁がしきい値よりも小さい場合は、第2の遅延、ゲイン計算回路393の処理は行なわずに、第1の遅延T₁を出力する。

【0097】処理を継続する場合は、第2の遅延、ゲイン計算回路393は、減算器394に出力に対して、式(5)及び式(6)のx_w(n)をe₁(n)におきかえて第2の遅延、ゲインを計算する。

【0098】第2のピッチ予測利得計算回路720は、式(35)においてx_w(n)をe₁(n)におきかえて第2のピッチ予測利得G₂を計算する。

【0099】判別回路730は、G₂をしきい値と判別し、あらかじめ定められたしきい値よりも大きいときは、第1の遅延と第2の遅延を出力する。あらかじめ定められたしきい値よりもG₂が小さい時は、第1の遅延のみを出力する。さらに、遅延の個数を端子399から出力する。

【0100】図7にもどって、音源量子化回路850は、まず、ピッチ抽出回路700から出力される遅延の個数を調べ、遅延の個数が2であれば、通常のビット数(B₁ビット)である第1の音源コードブック851を使用して音源信号を量子化するが、遅延の個数が1のときは、遅延を表すビット数と同一のビット数(B₂)の

【0092】図7は、第5の実施例の構成を示すブロック図である。図において、図1の構成と異なるのは、ピッチ抽出回路700、音源量子化回路850、第1の音源コードブック851、第2の音源コードブック852であるので、これらを説明する。

【0093】図8はピッチ抽出回路700の構成を示すブロック図である。

【0094】第1のピッチ予測利得計算回路710では、第1の遅延、ゲイン計算回路392で求めた遅延を用いて第1のピッチ予測利得を下記の数式22から求める。

【0095】

【数22】

第2の音源コードブック852を音源コードブック851と併用する。

【0101】図9は第6の実施例の構成を示すブロック図である。図3において、図7のピッチ抽出回路700を用いたものである。

【0102】音源量子化回路860は、ピッチ抽出回路700から遅延の個数を入力し、遅延の個数により、パルスの個数をM₁とM₂(M₁ < M₂)に切替え、さらに2種のビット数の異なる振幅コードブックを切替える。遅延の個数が2個の時はパルスの個数はM₁とし、第1の振幅コードブック861を用いる。遅延の個数が1のときはパルスの個数をM₂とし、第2の振幅コードブック862を用いる。

【0103】図10は、第7の実施例の構成を示すブロック図である。

【0104】図7において、図5のモード判別回路900を付加し、ピッチ抽出回路800は、あらかじめ定められたモードのときに、図7のピッチ抽出回路700と同一の動作を行なう。

【0105】図11は、第8の実施例の構成を示すブロック図である。図11は、図9において、図10に示したモード判別回路900、ピッチ抽出回路800を付加したものであるので、説明を省略する。

【0106】上述した実施例に限らず、種々の変形が可能である。

【0-1-0-7】モード情報を用いて、音源量子化回路や、ゲインコードブックを切替える構成とすることもできる。

【0108】音源コードブックを用いる場合、式(15)で示した歪みの小さい順に、複数のコードベクトルを選択し、ゲイン量子化回路でゲインを量子化しながら、式(18)を最小化する音源コードベクトルとゲインコードベクトルの組合せを選択しても良い。

【0109】また、パルス列で音源を表す場合、パルスの振幅を量子化するさいに、パルスの位置を複数セット求め、これらの各々に対して振幅コードブックを探索し、式(26)を最大化する組合せを選択してもよい。また、これらの組合せを複数種類ゲイン量子化回路に出力し、ゲイン量子化しながら、式(29)を最小化するような位置、振幅コードベクトル、ゲインコードベクトルの組合せを選択しても良い。

【0110】振幅コードブック540のかわりに、あらかじめ定められたビット数の極性コードブックを用いても良い。

【0111】ピッチ抽出回路で求めた複数の遅延は、差分符号化することにより量子化ビット数を削減できる。

【0112】

【発明の効果】以上説明したように、本発明によれば、入力信号から複数の遅延を求めた後に音源信号を計算することで、複数話者からなる音声や、複数の楽器から構成される音楽信号に対して、従来よりも良好な音質が得られるという効果がある。

【0113】さらに、遅延を求めながらピッチ予測利得を求め、ピッチ予測利得があらかじめ定められた条件を満たすかどうかを判別することで遅延の個数を可変にしているの、入力信号の特徴に応じて遅延の個数を適切に選ぶことができ、入力信号を良好に符号化することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施例による信号符号化装置のブロック図である。

【図2】図1の信号符号化装置のピッチ抽出回路390のブロック図である。

【図3】本発明の第2の実施例による信号符号化装置のブロック図である。

【図4】図3の信号符号化装置の音源量子化回路500のブロック図である。

【図5】本発明の第3の実施例による信号符号化装置の

ブロック図である。

【図6】本発明の第4の実施例による信号符号化装置のブロック図である。

【図7】本発明の第5の実施例による信号符号化装置のブロック図である。

【図8】図7の信号符号装置のピッチ抽出回路700のブロック図である。

【図9】本発明の第6の実施例による信号符号化装置のブロック図である。

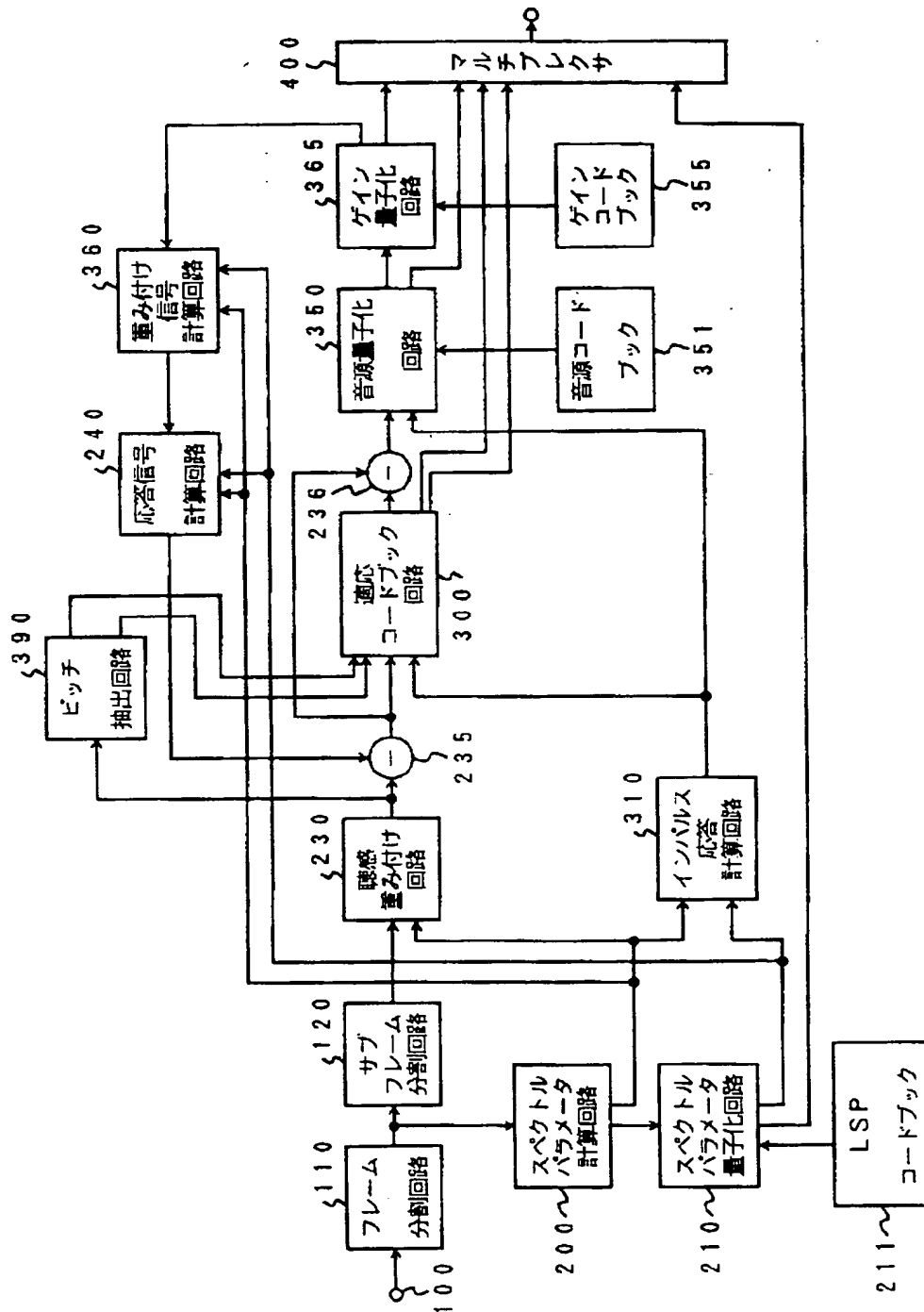
【図10】本発明の第7の実施例による信号符号化装置のブロック図である。

【図11】本発明の第8の実施例による信号符号化装置のブロック図である。

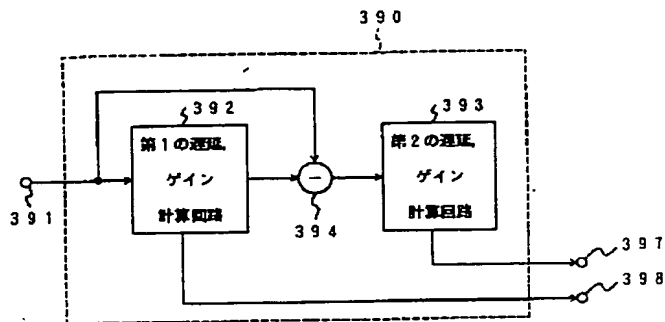
【符号の説明】

110	フレーム分割回路
120	サブフレーム分割回路
200	スペクトルパラメータ計算回路
210	スペクトルパラメータ量子化回路
230	聴感重み付け回路
235、236	減算回路
240	応答信号計算回路
300	適応コードブック回路
310	インパルス応答計算回路
350、500、850、860	音源量子化回路
351	音源コードブック
355、560	ゲインコードブック
365、550	ゲイン量子化回路
390、600、700、800	ピッチ抽出回路
392	第1の遅延計算回路
393	第2の遅延計算回路
400	マルチプレクサ
510	相関係数計算回路
520	位置計算回路
530	振幅量子化回路
540	振幅コードブック
710	第1のピッチ予測利得計算回路
720	第2のピッチ予測利得計算回路
730	判別回路
851	第1の音源コードブック、
852	第2の音源コードブック、
861	第1の振幅コードブック
862	第2の振幅コードブック
900	モード判別回路

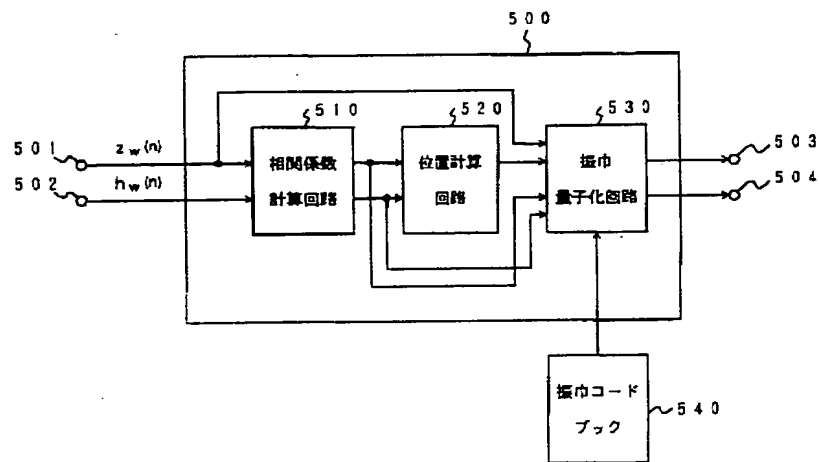
【図1】



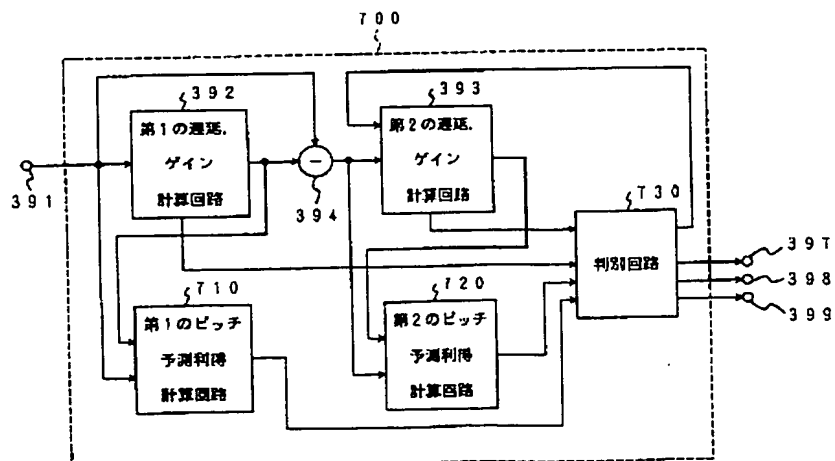
【図2】



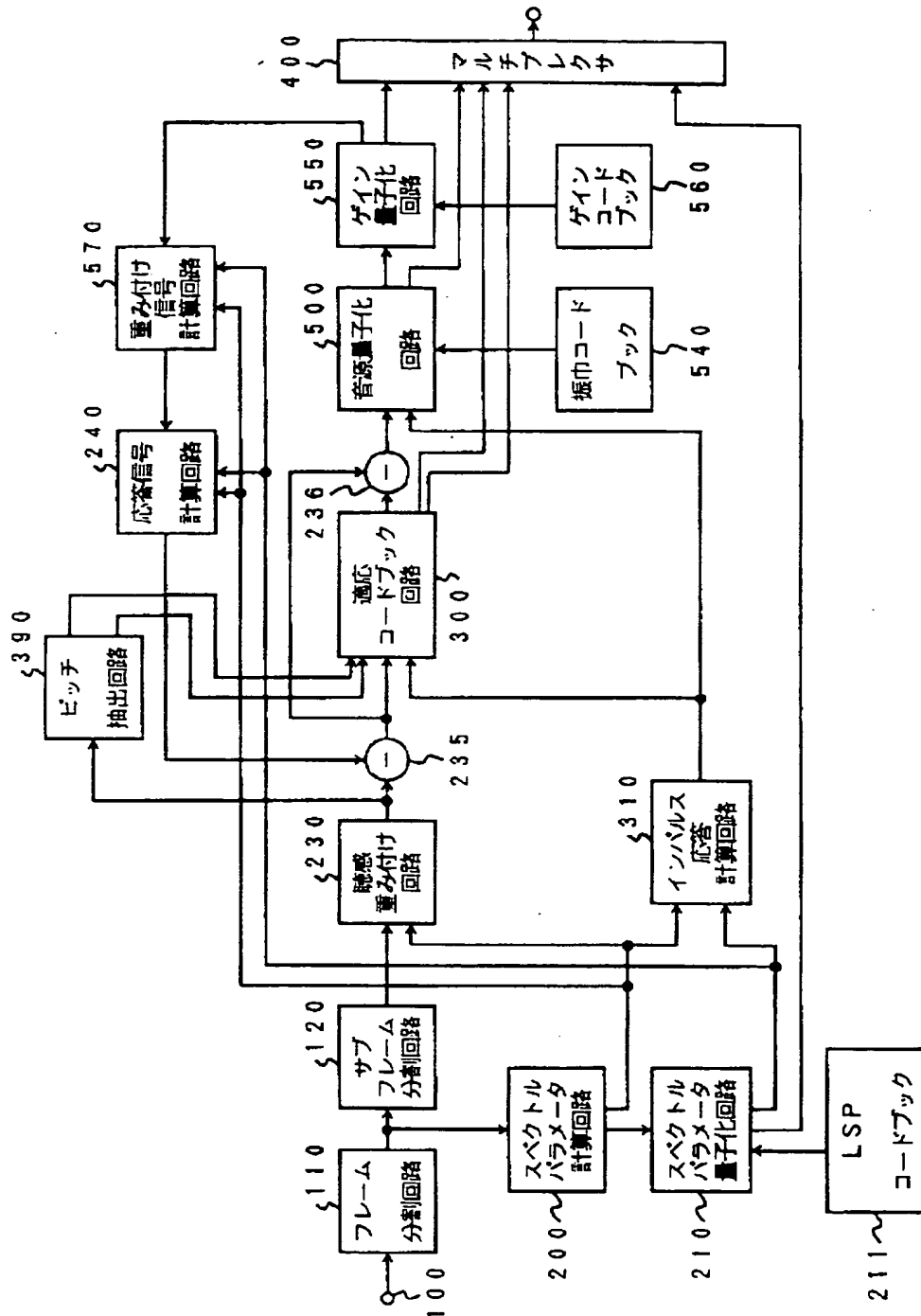
【図4】



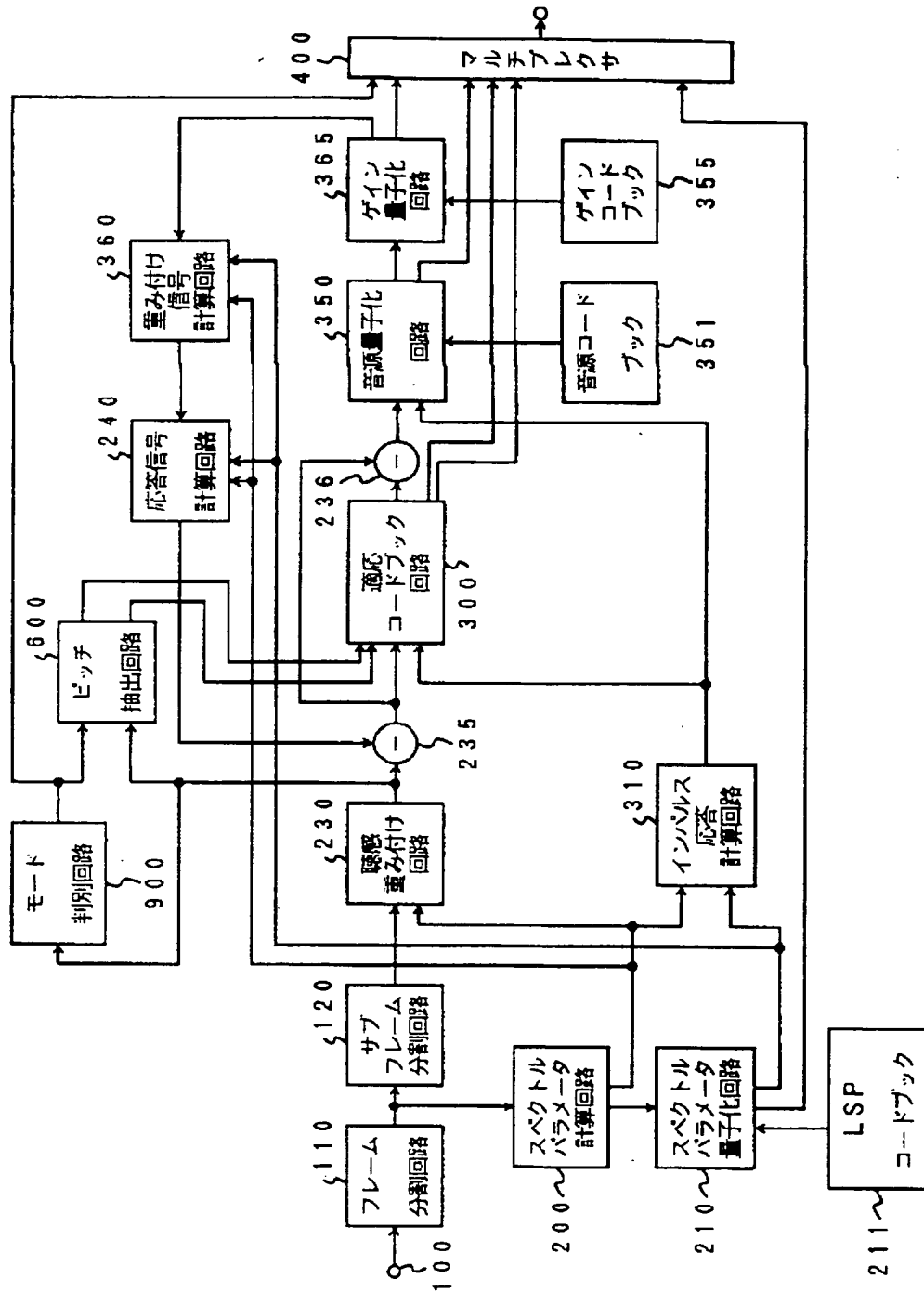
【図8】



【図3】



【図5】



[illegible]

【図 7】

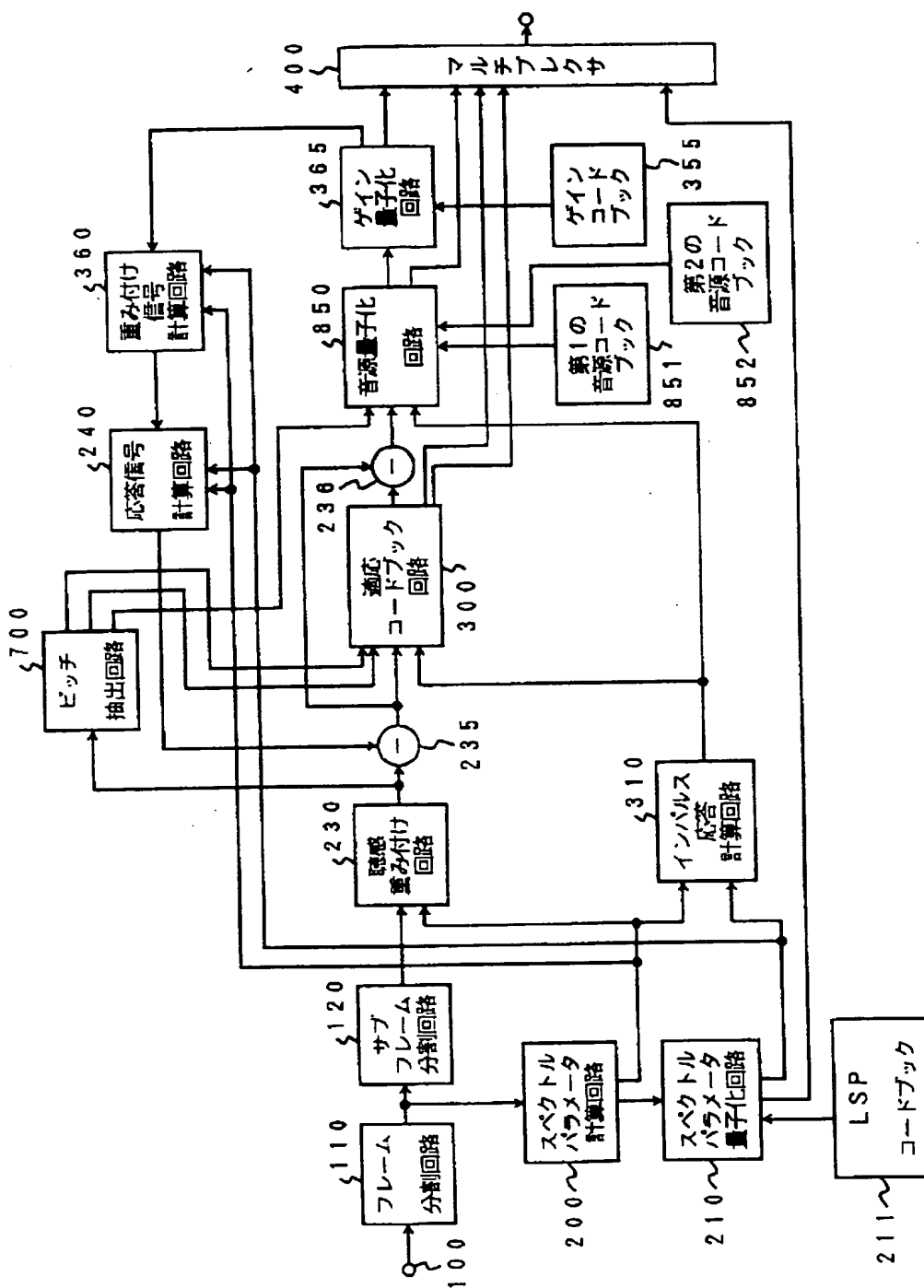
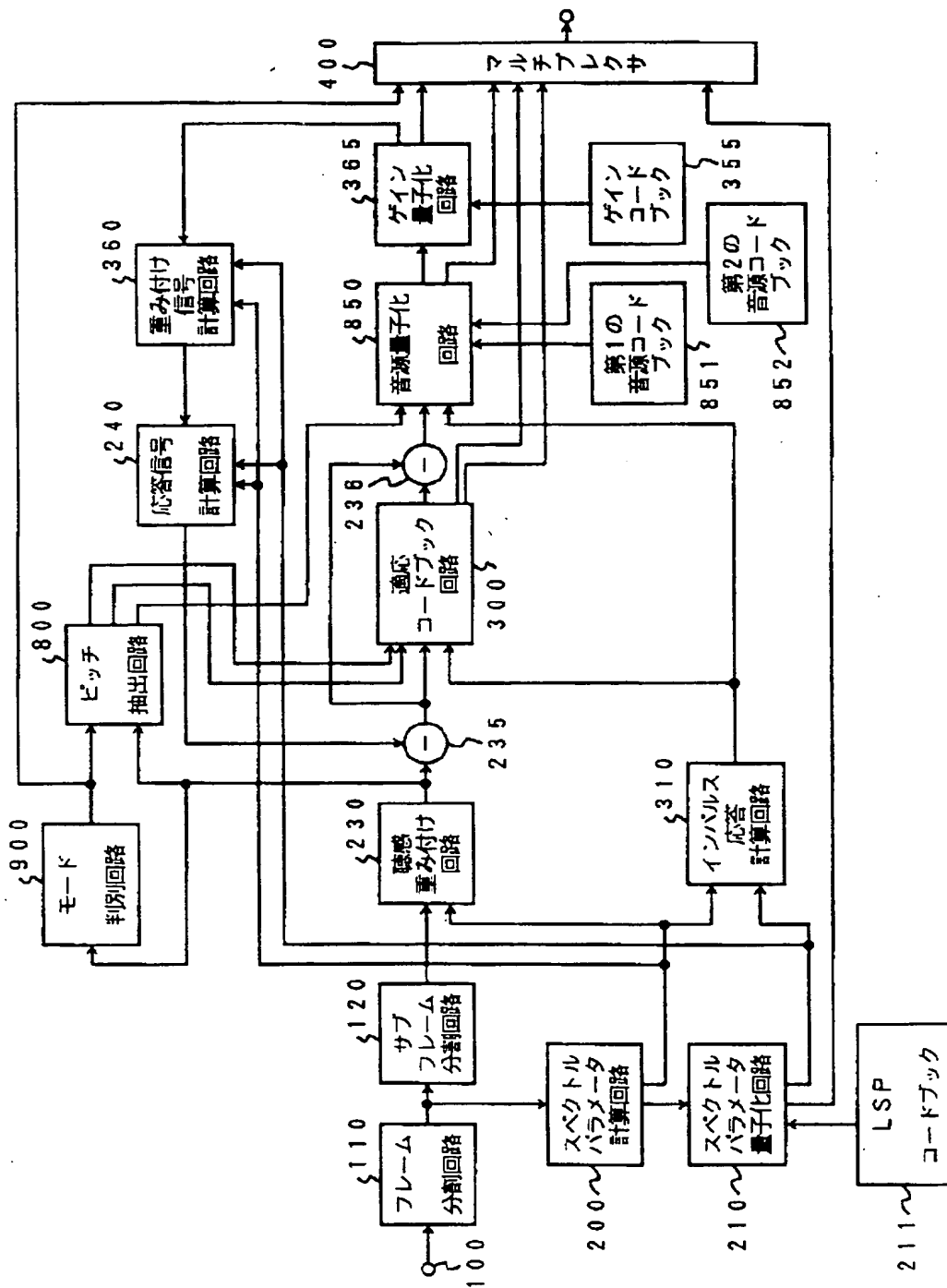


Fig. 1 is a block diagram of a speech processing system. The system includes the following components and their interconnections:

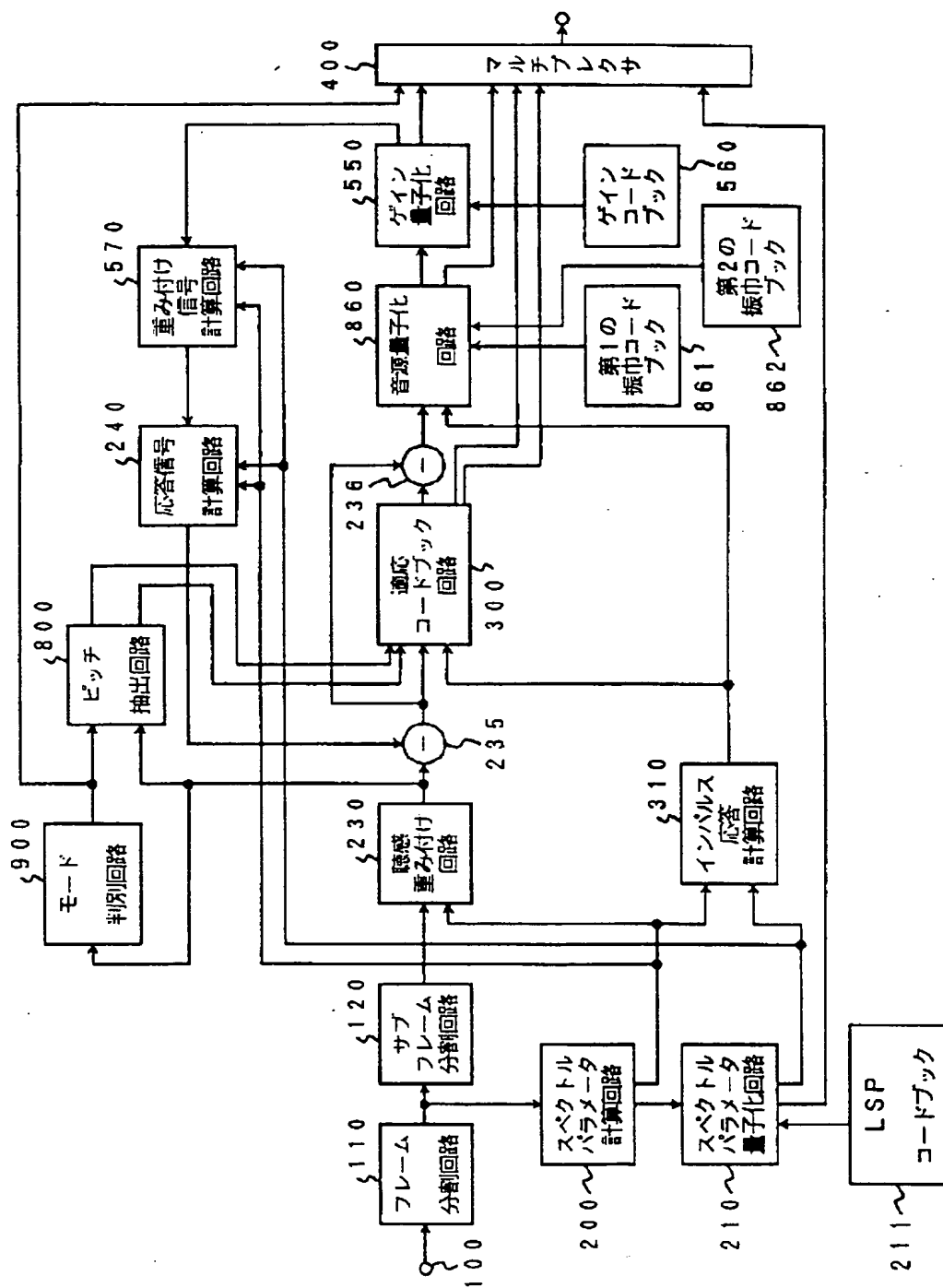
- Frame Division Circuit (100)**: Receives an input signal and outputs a frame signal.
- Sub-frame Division Circuit (120)**: Receives the frame signal and outputs a sub-frame signal.
- Spectrum Parameter Calculation Circuit (200)**: Receives the sub-frame signal and outputs spectrum parameters.
- Spectrum Parameter Quantization Circuit (210)**: Receives the spectrum parameters and outputs quantized spectrum parameters.
- LSP Codebook (211)**: Provides LSP codes to the spectrum parameter quantization circuit.
- Input Response Calculation Circuit (310)**: Receives the quantized spectrum parameters and outputs an input response signal.
- Adaptive Codebook Circuit (300)**: Receives the input response signal and outputs an adaptive codebook signal.
- Pitch Extraction Circuit (700)**: Receives the adaptive codebook signal and outputs a pitch signal.
- Response Signal Calculation Circuit (240)**: Receives the pitch signal and outputs a response signal.
- Weighted Signal Calculation Circuit (570)**: Receives the response signal and outputs a weighted signal.
- Audio Quantization Circuit (860)**: Receives the weighted signal and outputs a quantized audio signal.
- Gain Quantization Circuit (550)**: Receives the quantized audio signal and outputs a quantized gain signal.
- Gain Codebook (560)**: Provides gain codes to the gain quantization circuit.
- First Excitation Codebook (861)**: Provides first excitation codes to the audio quantization circuit.
- Second Excitation Codebook (862)**: Provides second excitation codes to the audio quantization circuit.

The diagram shows the flow of signals between these components, including feedback loops and various input/output points.

【図10】



【図 11】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☒ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.